Image texture features predict renal function decline in patients with autosomal dominant polycystic kidney disease

Timothy L. Kline[1], Panagiotis Korfiatis[1], Marie E. Edwards[2], Kyongtae T. Bae[3], Alan Yu[4], Arlene B. Chapman[5], Michal Mrug[6], Jared J. Grantham[4], Douglas Landsittel[3], William M. Bennett[7], Bernard F. King[1], Peter C. Harris[2], Vicente E. Torres[2], Bradley J. Erickson[1] and the CRISP Investigators

[1]Department of Radiology, Mayo Clinic College of Medicine, Rochester, Minnesota, USA; [2]Division of Nephrology and Hypertension, Mayo Clinic College of Medicine, Rochester, Minnesota, USA; [3]University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA; [4]The Kidney Institute, Department of Internal Medicine, Kansas University Medical Center, Kansas City, Kansas, USA; [5]Department of Medicine, University of Chicago, Chicago, Illinois, USA; [6]Division of Nephrology, University of Alabama and Department of Veterans Affairs Medical Center, Birmingham, Alabama, USA; and [7]Legacy Transplant Services, Legacy Good Samaritan Hospital, Portland, Oregon, USA

Magnetic resonance imaging (MRI) examinations provide high-resolution information about the anatomic structure of the kidneys and are used to measure total kidney volume (TKV) in patients with Autosomal Dominant Polycystic Kidney Disease (ADPKD). Height-adjusted TKV (HtTKV) has become the gold-standard imaging biomarker for ADPKD progression at early stages of the disease when estimated glomerular filtration rate (eGFR) is still normal. However, HtTKV does not take advantage of the wealth of information provided by MRI. Here we tested whether image texture features provide additional insights into the ADPKD kidney that may be used as complementary information to existing biomarkers. A retrospective cohort of 122 patients from the Consortium for Radiologic Imaging Studies of Polycystic Kidney Disease (CRISP) study was identified who had T2-weighted MRIs and eGFR values over 70 mL/min/1.73m^2 at the time of their baseline scan. We computed nine distinct image texture features for each patient. The ability of each feature to predict subsequent progression to CKD stage 3A, 3B, and 30% reduction in eGFR at eight-year follow-up was assessed. A multiple linear regression model was developed incorporating age, baseline eGFR, HtTKV, and three image texture features identified by stability feature selection (Entropy, Correlation, and Energy). Including texture in a multiple linear regression model (predicting percent change in eGFR) improved Pearson correlation coefficient from -0.51 (using age, eGFR, and HtTKV) to -0.70 (adding texture). Thus, texture analysis offers an approach to refine ADPKD prognosis and should be further explored for its utility in individualized clinical decision making and outcome prediction.

Kidney International (2017) ■, ■–■; http://dx.doi.org/10.1016/j.kint.2017.03.026

Correspondence: *Bradley J. Erickson, 200 First Street SW, Rochester, Minnesota 55905, USA. E-mail: bje@mayo.edu*

Received 12 July 2016; revised 10 March 2017; accepted 16 March 2017

KEYWORDS: gray-level co-occurrence matrix; magnetic resonance imaging; multiple linear regression; polycystic kidney disease; total kidney volume

Radiological imaging plays an important role in diagnosis and management of patients with autosomal dominant polycystic kidney disease (ADPKD).[1] Total kidney volume (TKV), measured by magnetic resonance imaging (MRI), computed tomography (CT), or ultrasound, predicts the decline of renal function[2–4] and is used to ascertain the effectiveness of treatment interventions.[5–7] Recently, the US Food and Drug Administration and the European Medicines Agency qualified TKV, in combination with patient age and estimated glomerular filtration rate (eGFR), as a prognostic imaging biomarker for ADPKD research and clinical trials. Particularly, height-adjusted TKV (HtTKV) is the strongest prognostic biomarker of renal function worsening and progression to end-stage renal disease (ESRD) in patients at an early stage of ADPKD when eGFR is within the normal range. This has led to the development of a number of automated approaches to accurately and reproducibly measure HtTKV.[8–11] Although HtTKV provides important information regarding kidney size, which is linked to overall disease progression, it does not take advantage of the wealth of information available within an MRI or CT image related to renal tissue structure.

There are a number of possible imaging biomarkers beyond HtTKV. One would be to measure the cystic load, including quantification of the number, size, and composition of renal cysts. This could lead to a deeper understanding of a patient's particular phenotype in terms of whether, for example, the kidneys are composed of a few large simple exophytic cysts or have numerous complex (e.g., proteinaceous) cysts. However, performing segmentation of renal cysts from radiological scans of ADPKD patients is extremely time-consuming and challenging, particularly in patients with a moderate to severe cyst burden.[12] Therefore, automated

approaches are being actively pursued. A number of studies have published cyst segmentation techniques. These techniques have included threshold-based approaches,[13,14] as well as more advanced, semi-automated techniques including shape or boundary detection.[15] In a majority of these studies, T2-weighted MRI scans are used due to the difference in signal intensity of fluid-filled cysts (high signal intensity) compared with renal parenchyma (low signal intensity).[16] However, these techniques typically rely on the assumption that all cysts have similar signal intensities. However, complex cysts often are darker, and therefore most often only simple fluid-filled cysts are delineated. Techniques that can differentiate not only cysts from parenchyma but also classify different types of cysts (simple, proteinaceous, or infected) will be necessary to accurately characterize renal cystic burden in more detail.

Another opportunity is through the use of quantitative MRI techniques.[17] These techniques could provide a deeper understanding of the composition of an ADPKD patient's renal tissue. For example, magnetization transfer imaging has been shown to identify fibrotic tissue within kidneys,[18–20] perfusion imaging (e.g., arterial spin labeling) can identify reduced perfusion territories,[21] and blood oxygen level–dependent imaging reflects regional renal oxygenation status.[22–24] Other techniques such as measurement of renal blood flow,[25–27] magnetic resonance elastography (to measure tissue stiffness),[28] sodium imaging,[29] diffusion-weighted imaging[30] with intravoxel incoherent motion processing,[31] and chemical exchange saturation transfer imaging[32] are additional imaging techniques currently being explored that could show utility in ADPKD management. However, the information regarding how these imaging findings relate to progression of the disease in humans may not be available for quite some time.

Advanced image processing techniques offer a unique opportunity to retrospectively explore the large number of imaging examinations available and correlate this information with known progression of the disease. Texture analysis is 1 technique that performs an ensemble of mathematical computations to greatly increase the information provided by conventional radiological images.[33] Image texture refers to the appearance, structure, and arrangement of different intensity levels within an image and has been used in organ segmentation tasks (e.g., segmentation of kidneys in ultrasound images[34] and CT images[35]), as well as clinical decision making (e.g., classifying breast tumors,[36] early diagnosis of multiple sclerosis,[37] and differentiating liver lesions[38]).

In this manuscript, we examined whether image textures can serve as imaging biomarkers in ADPKD. We hypothesize that texture-based imaging biomarkers provide quantifiable parameters of renal tissue structure that will better classify patients and improve individual prognosis compared with HtTKV.

A retrospective cohort from the Consortium for Radiologic Imaging Studies of Polycystic Kidney Disease (CRISP) study was identified, consisting of 122 patients who had T2-weighted MRIs, eGFR values >70 ml/min/1.73 m^2 at the time of their baseline scan, and typical presentations of the disease, according to the definition provided by Irazabal *et al.*[39] Based on their eGFR measurements performed 8 years later, the patients were split into 2 groups according to whether they had or had not reached each of the following: chronic kidney disease (CKD) stage 3A, stage 3B, and a 30% reduction in eGFR (a potentially acceptable surrogate endpoint for the development of ESRD in some circumstances).[40]

RESULTS

Patient cohort

In order to explore the potential of texture features as candidate imaging biomarkers, a retrospective cohort of patients from the CRISP study was identified with T2-weighted MRIs, eGFR values >70 ml/min/1.73 m^2 at the time of their baseline scan, and typical presentations of the disease ($N = 122$) based on the description provided by Irazabal *et al.*[39] Based on eGFR measurements performed 8 years later, the patients were split into 2 groups according to whether they had ($n = 44$) or had not ($n = 78$) reached CKD stage 3A (eGFR <60 ml/min/1.73 m^2), whether they had ($n = 22$) or had not ($n = 100$) reached CKD stage 3B (eGFR <45 ml/min/1.73 m^2), and whether a 30% decrease of eGFR had ($n = 47$) or had not ($n = 75$) occurred.

A wide range of disease phenotypes were observed in this study's cohort. Shown in Figure 1 are examples of 6 patients with different renal cyst appearances along with age, sex, HtTKV, and genetic mutation information. Although some of these differences are obvious and one might suspect clinically that the disease severity would be different, no quantifiable parameters (i.e., imaging biomarkers) have yet been developed to characterize these differences.

Texture analysis

Image texture features offer a promising approach to provide additional insights into PKD phenotypes. Figure 2 depicts how image texture features convey differences in tissue heterogeneity. In particular, it demonstrates how similar signal intensity regions in the MR image have drastically different textural depictions. Figure 3 shows the 9 image texture features calculated in this study, based on a single middle coronal slice from the same patient. Each image texture feature calculates distinct properties, reflecting differences in kidney phenotype.

Predictive power of biomarkers

At baseline, age and eGFR had relatively little predictive value of subsequent renal function decline, whereas HtTKV had a much better predictive value, as presented in Table 1 and shown in the receiver operating characteristic curves of Figure 4. For the 3 comparisons (progression to CKD stage 3A, progression to CKD stage 3B, and 30% change in eGFR), HtTKV had strong predictive power for progression to CKD stage 3A and 3B ($A_z = 0.81$ and 0.84, respectively), while it

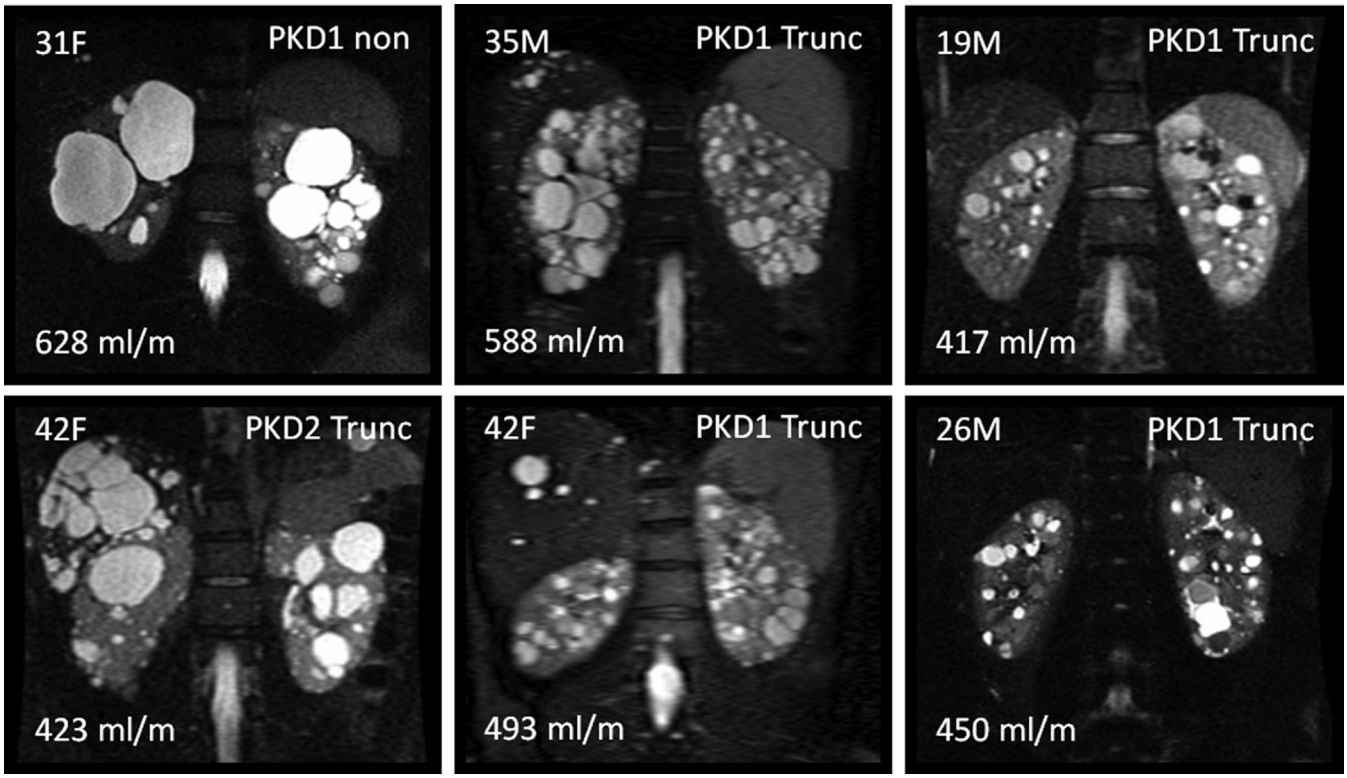

Figure 1 | Examples highlighting phenotypic differences of 6 different autosomal dominant polycystic kidney disease (ADPKD) patients from this study. Shown in each panel is the age and sex of the patient in the upper left, the genetic mutation in the upper right, and the height-adjusted total kidney volume (TKV) measurement in the bottom left. Although visually many differences are apparent, no quantifiable image-based morphological parameters have yet been utilized beyond TKV and the cystic burden of kidneys to characterize ADPKD phenotypes. To optimize viewing of this image, please see the online version of this article at www.kidney-international.org.

was less informative for prediction of a 30% or more change in eGFR ($A_z = 0.73$). Shown in Figure 5 are the regression analysis results, comparing baseline age, eGFR, and HtTKV to percentage change in eGFR at 8-year follow-up. The Pearson correlation coefficient (r) was –0.07 ($P = 0.45$) for age, –0.03 ($P = 0.74$) for baseline eGFR, and –0.48 ($P < 0.01$) for HtTKV. There were a number of patients whose prognosis was not accurately determined using these more traditional biomarkers by themselves. A multiple linear regression model incorporating age, baseline eGFR, and HtTKV performed

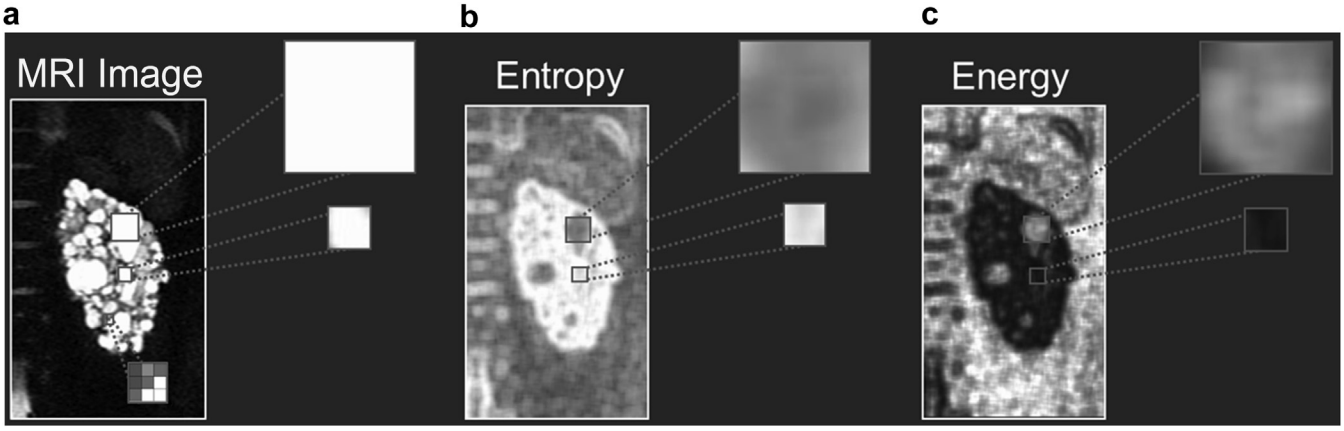

Figure 2 | Texture feature analysis is an image processing technique that can elucidate phenotypic differences of the kidneys. (**a**) T2-weighted MRI of the left kidney of 1 of the patients in this study. If we consider 2 different cystic regions (insets), the information (in terms of image intensity) is similar. However, in the 2 example texture images of (**b**) entropy and (**c**) energy, information regarding size and position is ingrained in the image. Notice how the 2 cyst regions look very different in the texture images, and thus quantifiable values can be extracted that convey tissue structural information such as cyst size and number. Other structural tissue changes (e.g., fibrotic tissue or cyst classification) could likely be conveyed within the texture images. To optimize viewing of this image, please see the online version of this article at www.kidney-international.org.

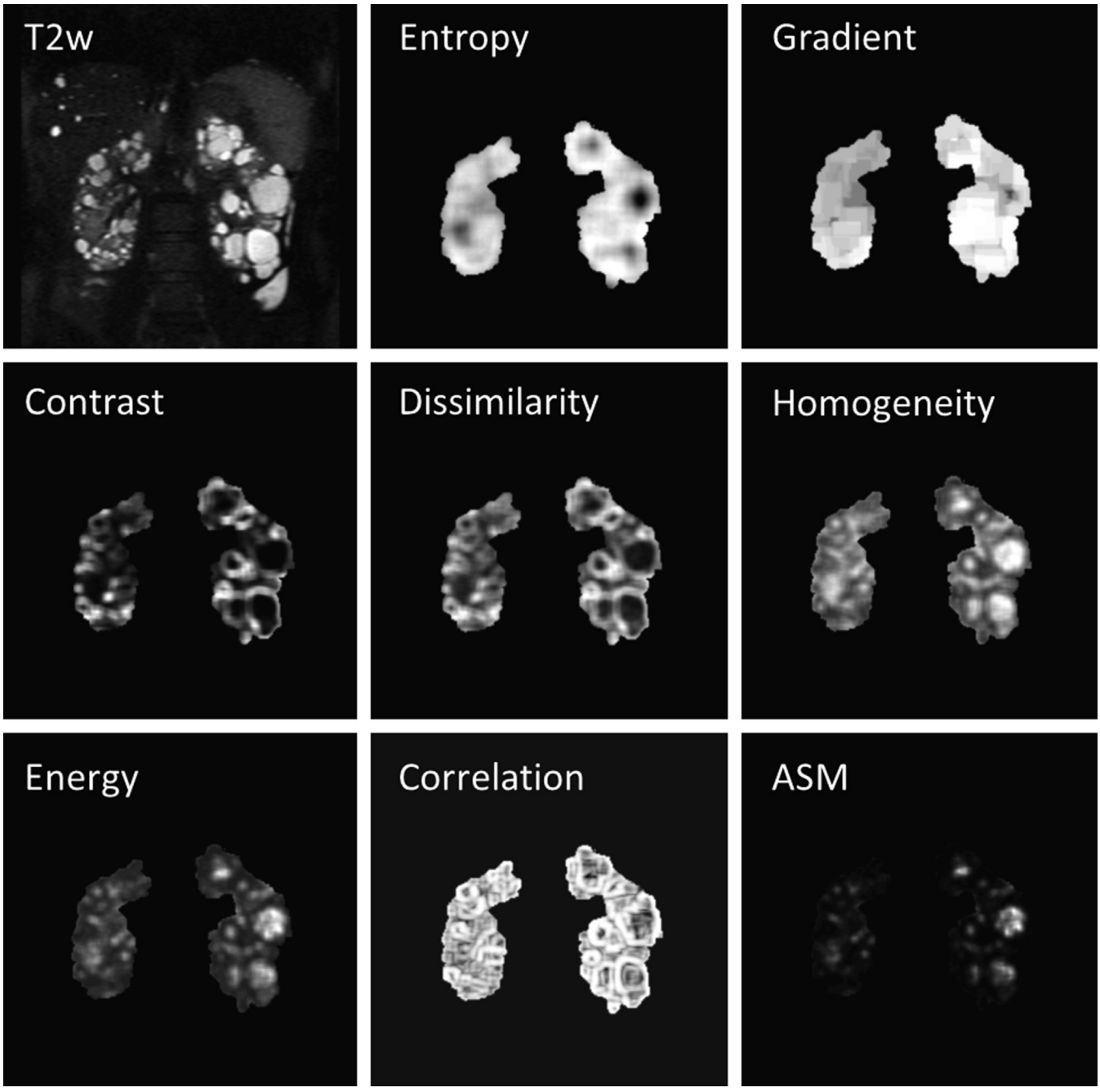

Figure 3 | Examples of the 9 derived imaging texture features for a single patient. From top left, by row then column: T2-weighted magnetic resonance image from which grayscale values are extracted and texture analysis is performed, first-order entropy, first-order gradient, second-order gray level co-occurrence matrix (GLCM) contrast, second-order GLCM dissimilarity, second-order homogeneity, second-order GLCM energy, second-order GLCM correlation, and second-order GLCM angular second moment (ASM). *Entropy* measures the degree of disorder within the kidney. Kidneys with seemingly random cyst distributions will have a higher entropy value. Note how large cystic regions appear dark, whereas in regions of many small cysts a high entropy value is calculated. *Gradient* is a measure of grayscale changes. Kidneys with many small cysts will have more detectable edges and larger local changes in grayscale. *Contrast* measures the local variations in the GLCM. Low values mean that the gray levels are similar throughout the kidneys. *Dissimilarity* measures differences in GLCM elements, which relates to a measure of renal tissue heterogeneity. *Homogeneity* measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. Note how both large cystic regions as well as regions with no cystic burden appear bright. *Energy* provides the sum of squared elements in the GLCM and is a measure of tissue uniformity. It is closely related to the inverse of entropy. *Correlation* measures grayscale value dependence of kidney voxels (the joint probability occurrence of specified pixel pairs). *Angular second moment* (ASM) is also a measure of tissue homogeneity. To optimize viewing of this image, please see the online version of this article at www.kidney-international.org.

Table 1 | Summary statistics for age, eGFR, HtTKV, and the 3 image texture features identified by the stability feature selection approach

Progression to CKD stage 3A							
Biomarker	**>60**	**<60**	***P* value**	**Sensitivity**	**Specificity**	**A_z**	**Threshold**
Age	30.0 ± 9.0	34.6 ± 5.5	0.013	0.89	0.41	0.64 [0.58–0.72]	26.5
eGFR	100.0 ± 16.0	81.3 ± 5.6	<0.01[a]	0.68	0.83	0.77 [0.62–0.91]	94.2
HtTKV	403 ± 118	724 ± 239	<0.01[a]	0.82	0.67	0.81 [0.75–0.87]	457
Entropy	8.80 ± 0.23	9.36 ± 0.17	<0.01[a]	0.84	0.96	0.93 [0.88–0.97]	9.2
Correlation	0.51 ± 0.05	0.55 ± 0.02	<0.01[a]	0.80	0.63	0.72 [0.64–0.79]	0.53
Energy	1.17 ± 0.35	1.82 ± 0.34	<0.01[a]	0.57	0.90	0.80 [0.74–0.87]	1.74
Progression to CKD stage 3B							
Biomarker	**>45**	**<45**	***P* value**	**Sensitivity**	**Specificity**	**A_z**	**Threshold**
Age	31.7 ± 7.5	34.6 ± 5.9	0.17	0.96	0.28	0.59 [0.50–0.70]	24.6
eGFR	96.3 ± 14.2	80.8 ± 4.4	<0.01[a]	0.52	1.00	0.75 [0.70–0.81]	94.2
HtTKV	428 ± 143	896 ± 324	<0.01[a]	0.73	0.83	0.82 [0.73–0.90]	700
Entropy	8.92 ± 0.28	9.38 ± 0.17	<0.01[a]	0.87	0.76	0.86 [0.78–0.91]	9.1
Correlation	0.52 ± 0.05	0.57 ± 0.02	<0.01[a]	1.00	0.58	0.79 [0.71–0.85]	0.53
Energy	1.25 ± 0.34	1.86 ± 0.23	<0.01[a]	1.00	0.51	0.80 [0.74–0.87]	1.26
Percentage change in eGFR							
Biomarker	**<30%**	**>30%**	***P* value**	**Sensitivity**	**Specificity**	**A_z**	**Threshold**
Age	31.8 ± 7.4	32.1 ± 7.3	0.469	0.77	0.35	0.54 [0.47–0.62]	26.5
eGFR	94 2 ± 12.7	87.4 ± 9.3	0.350	0.45	0.77	0.55 [0.45–0.63]	97.7
HtTKV	415 ± 129	610 ± 215	<0.01[a]	0.81	0.53	0.73 [0.64–0.80]	423
Entropy	8.83 ± 0.24	9.28 ± 0.24	<0.01[a]	0.70	0.85	0.82 [0.73–0.88]	9.1
Correlation	0.51 ± 0.05	0.55 ± 0.02	<0.01[a]	0.77	0.63	0.69 [0.62–0.78]	0.53
Energy	1.18 ± 0.34	1.74 ± 0.35	<0.01[a]	0.49	0.91	0.75 [0.69–0.83]	1.81

CKD, chronic kidney diseae; eGFR, estimated glomerular filtration rate; HtTKV, height-adjusted total kidney volume.
Statistics are given for progression to CKD stage 3A, stage 3B, and 30% cut-off threshold for percentage change in eGFR. Each group has the median and median absolute deviation listed between groups, *P* value calculated from Wilcoxon rank-sum test, sensitivity, specificity, area under receiver operating characteristic (ROC) curve (A_z) with 95% confidence interval, and the cutoff threshold used for ROC analysis.
[a]$P < 0.01$ was deemed significant.

relatively well at distinguishing between the patient groupings (Figure 4, model labeled "Traditional"), with $r = -0.51$ ($P < 0.01$), as shown in Figure 5d.

We used stability of texture features as the basis for selecting 3 texture metrics (median of entropy with window = 31, 25th percentile of correlation with window = 7, and skewness of energy with window = 7). Focusing on stability rather than just the best performing features for this cohort should enable better generalization to other patient cohorts. The correlation between the 3 traditional biomarkers (age, eGFR, and HtTKV) and the 3 image texture features is shown in Figure 4d. We found that these 3 texture features correlated well with a 30% or more decrease in eGFR and differentiated the patients who progressed to either CKD stage 3A or 3B (Table 1 and Figure 6). Entropy had the strongest predictive power of the 3, with $A_z = 0.93$ for progression to CKD stage 3A, $A_z = 0.86$ for progression to CKD stage 3B, and $A_z = 0.82$ for predicting a 30% or more reduction in eGFR.

Texture biomarkers also had a strong correlation with percentage change in eGFR as shown in Figure 6. The Pearson's *r* for entropy texture was –0.52 ($P < 0.01$), for correlation texture was –0.43 ($P < 0.01$), and for energy texture was –0.52 ($P < 0.01$). The 3 image texture features were included in a multiple linear regression model in order to ascertain the added value over traditional image-derived modeling (age + eGFR + HtTKV). Incorporating the texture features improved prediction based on correlation coefficient, obtaining an $r = -0.70$ ($P < 0.01$) as shown in Figure 6d. Also, summary statistics for the 2 models are shown in Table 2.

Reproducibility of texture features

Texture-derived imaging biomarkers were highly reproducible. In order to determine the reproducibility of the image-derived texture features, 30 cases were reanalyzed by a second observer. The 2 major sources for variability were segmentation of the kidneys and the measurement of cerebral spinal fluid (CSF), which is used to linearly normalize the MR scans. For the mean ± SD comparison of the 2 segmentations, the Jaccard similarity metric was 0.97 ± 0.01, precision was 0.98 ± 0.01, and percentage volume error was 0.95% ± 1.14%. For measurement of CSF, the percentage difference was 3.2% ± 5.0%. More importantly, the impact on texture measures was small: the percentage difference for entropy was 0.10% ± 0.23%, for correlation was 0.91% ± 3.2%, and for energy was 0.13% ± 0.35%.

DISCUSSION

This current study presents the first report of texture analysis for ADPKD assessment. A cohort of patients (from the CRISP

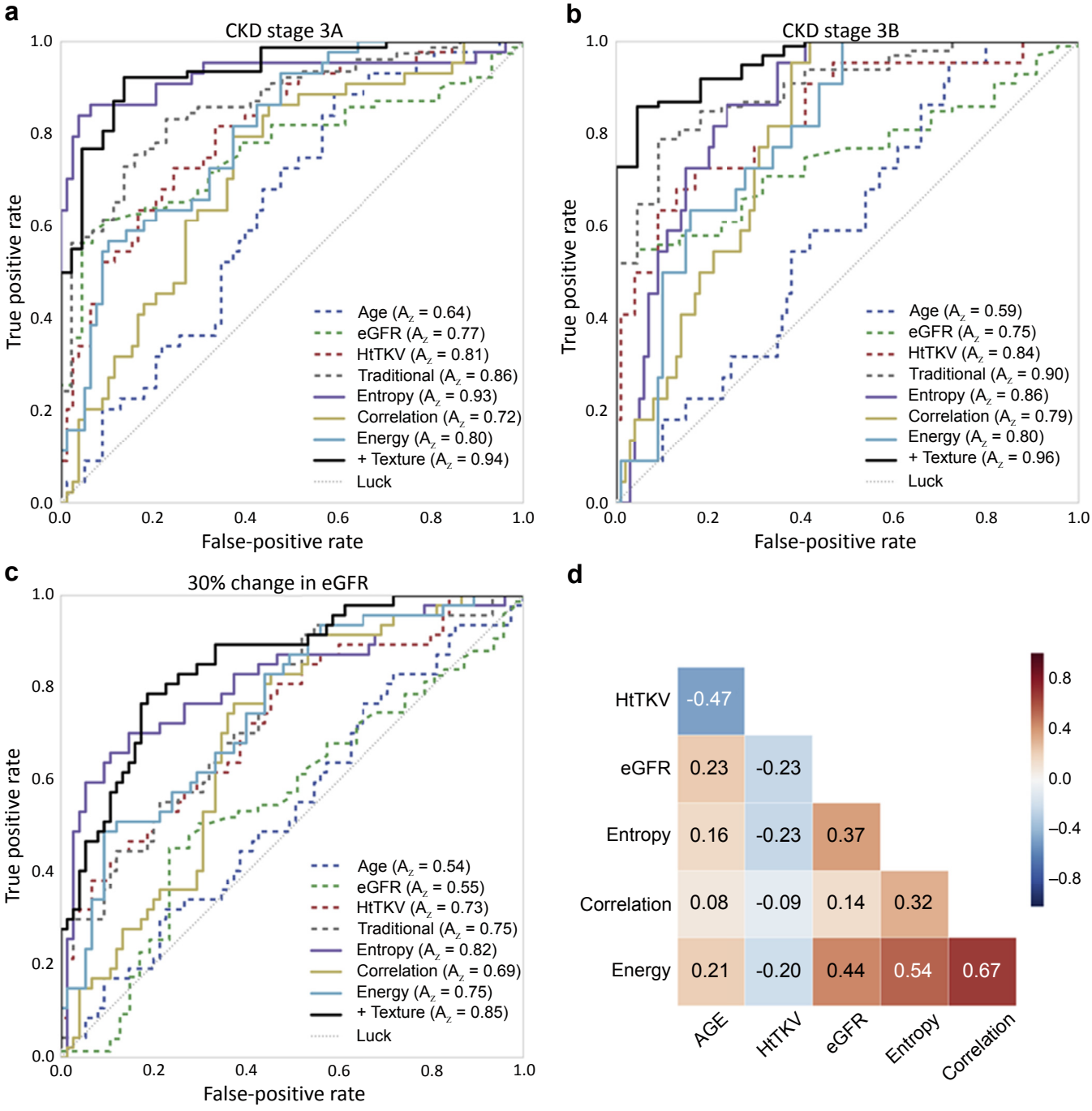

Figure 4 | Individual biomarkers show strong predictive power for progression to chronic kidney disease (CKD) stage 3A, CKD stage 3B, and 30% change in estimated glomerular filtration rate (eGFR) in 8-year follow-up examinations. (**a**) Receiver operating characteristic (ROC) curves for predicting progression to CKD stage 3A. (**b**) ROC curves for predicting progression to CKD stage 3B. (**c**) ROC curves for prediction of a 30% or more decrease in eGFR. The individual biomarkers include age, eGFR, height-adjusted total kidney volume (HtTKV), entropy, correlation, and energy. (**d**) Correlation matrix showing how the individual features relate to one another. These biomarkers were included in a multiple linear regression model in order to ascertain the added value over traditional image-derived modeling. The traditional model used age, eGFR, and HtTKV, while the texture model also included entropy, correlation, and energy. Compared with the traditional model, the addition of texture improved the predictive power for CKD stage 3A from $A_z = 0.86$ to $A_z = 0.94$, for CKD stage 3B from $A_z = 0.90$ to $A_z = 0.96$, and for a 30% or more change in eGFR from $A_z = 0.75$ to $A_z = 0.85$. Coloring in ROC curves corresponds to colors in Figures 5 and 6, and dashed biomarkers reflect traditional modeling, whereas solid lines reflect texture-based modeling.

study) with normal renal function at baseline and typical presentations of the disease was identified. We assessed the value of various biomarkers to predict progression to CKD stage 3A and 3B 8 years from baseline, as well as percentage change in eGFR. MRIs were preprocessed and standardized, and kidneys were segmented. Texture features were extracted

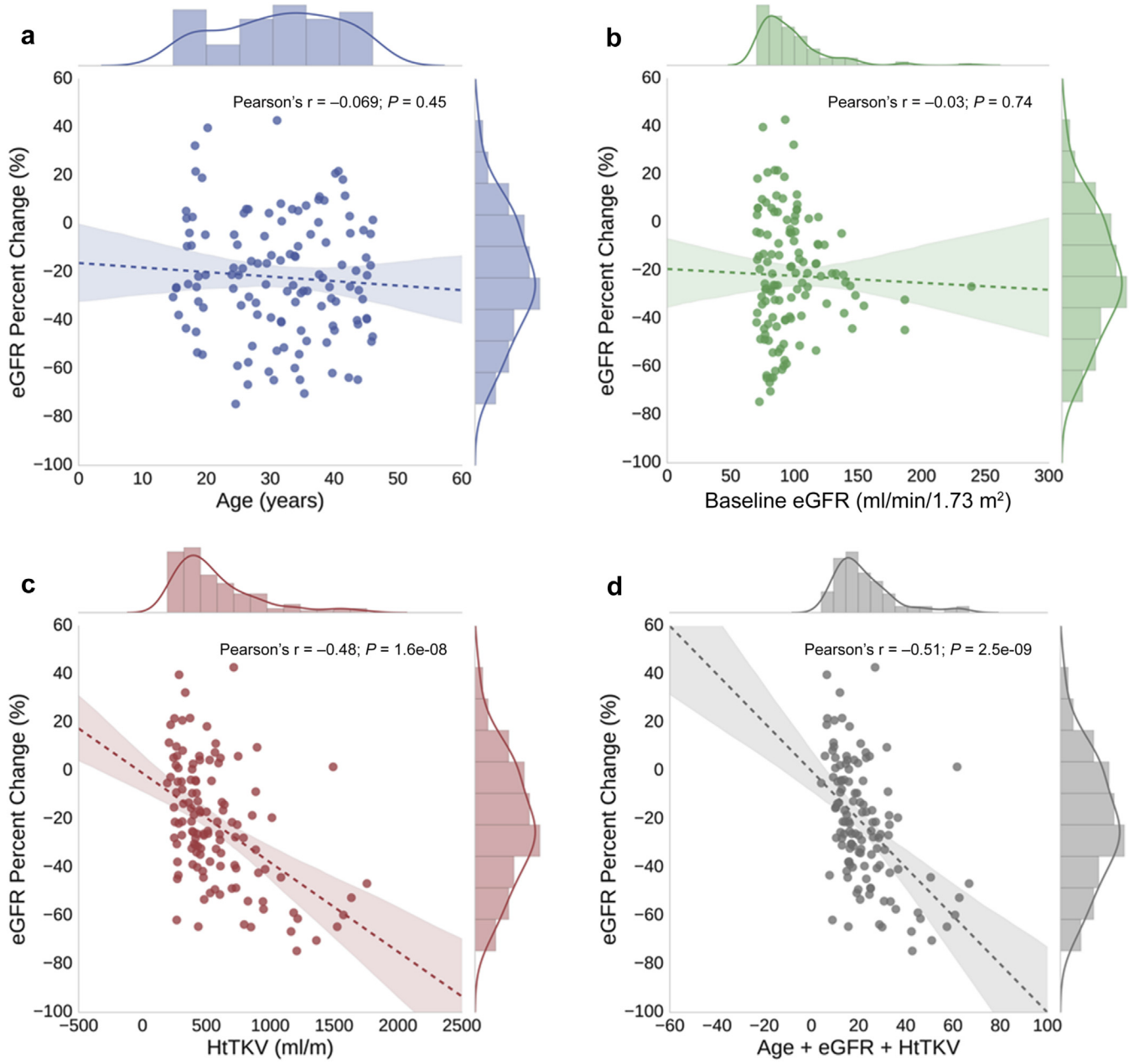

Figure 5 | Regression analysis results, comparing baseline traditional biomarkers (age, baseline estimated glomerular filtration rate [eGFR], and height-adjusted total kidney volume [HtTKV]) with percentage change in eGFR at the 8-year follow-up. (**a**) Age and (**b**) baseline eGFR have very low correlation, whereas (**c**) HtTKV has a fairly high correlation, with a subsequent percentage change in eGFR. (**d**) A multiple linear regression model incorporating these biomarkers obtained a Pearson's $r = -0.51$. Coloring corresponds to that used for each feature in Figure 4.

from the images and evaluated for their individual performance in distinguishing rapidly versus slowly progressing groups. Lastly, a multiple linear regression model was developed that demonstrated the added value of texture image analysis over current models.

Individual texture features were found to predict subsequent renal function decline. These features characterized factors such as the degree of disorder within the kidney, differences in cyst size and number, and whether the kidneys were composed of many similar-appearing regions to differentiate the patients. The combination of the features identified by stability selection utilizing the multiple linear regression model system allowed for the greatest prediction ability for subsequent renal function decline. It is likely that the combination of texture feature analysis with other non–image-based biomarkers will improve personalized clinical decision making in ADPKD. We believe that these texture metrics can function as a complementary tool to existing classification methods, such as the HtTKV-based prognostic model developed by Irazabal *et al.*[39] and the proPKD score

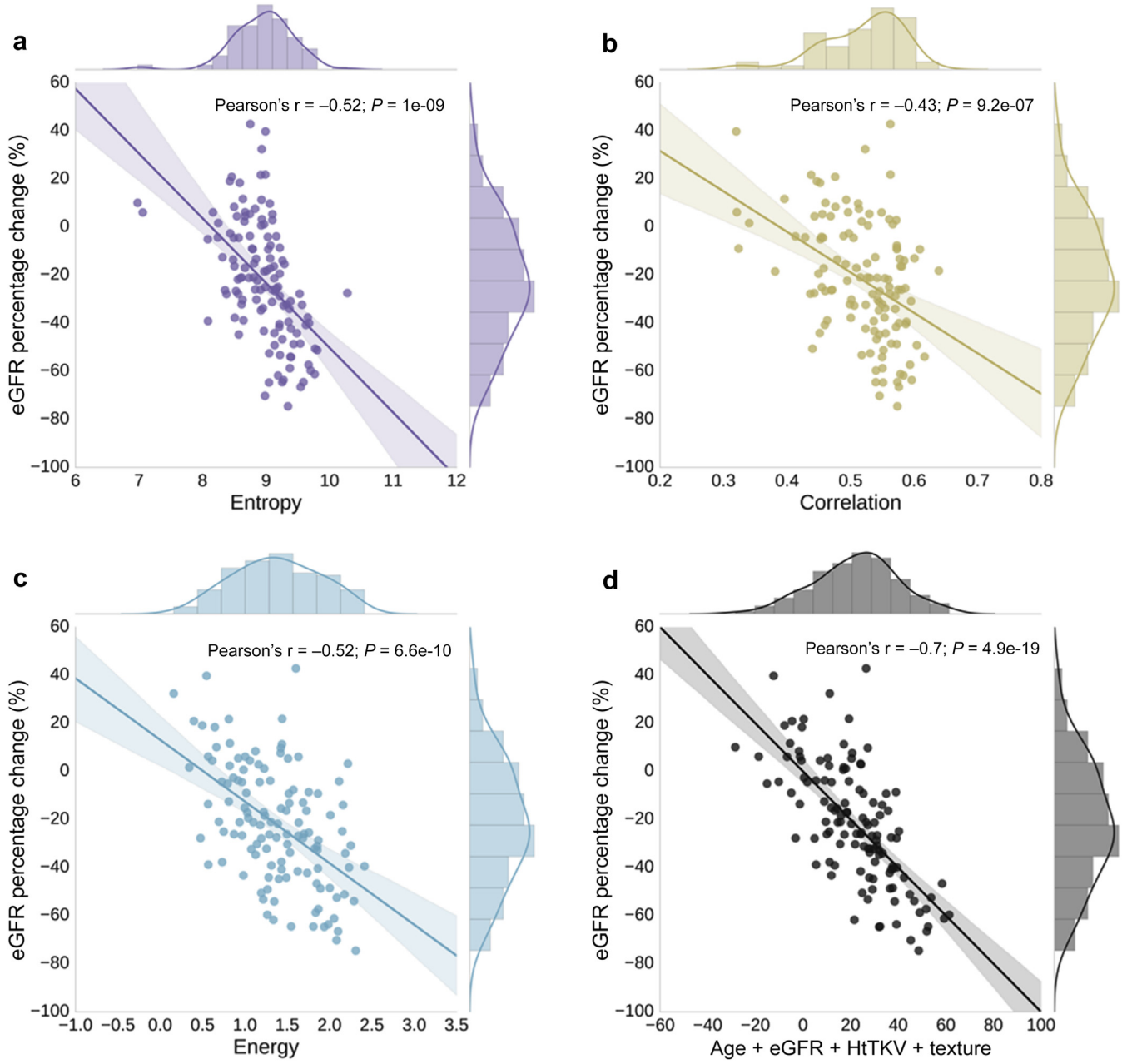

Figure 6 | Individual texture biomarkers show a strong correlation with the subsequent percentage change in estimated glomerular filtration rate (eGFR). Shown here are the regression analysis results, comparing the texture biomarkers (**a**) entropy, (**b**) correlation, and (**c**) energy with the percentage change in eGFR at the 8-year follow-up. (**d**) A multiple linear regression model incorporating traditional biomarkers and texture features. This multiple linear regression model incorporating image texture features helped further refine the prediction of subsequent renal function decline, improving Pearson's *r* from –0.51 (traditional biomarkers alone, Figure 5d) to –0.70 (addition of image texture features). Coloring corresponds to that used for each feature in Figure 4.

developed by Cornec-Le Gall *et al.*[41] In particular, based on these results (Table 2), there is a 23% narrowing of the prediction intervals when including the MR-derived texture features. As an example, using the traditional model, a prediction of a 30% decrease in eGFR has a 95% confidence interval with a fairly wide range (16.5%–43.5%), whereas including texture has a 95% confidence interval, which is much more precise (26.5%–33.5%). Thus, incorporating image texture into current prediction models greatly improves disease prognosis precision.

In addition, this approach of texture analysis may be preferable to other imaging methods including new quantitative MRI techniques, as they may be more challenging to put into practice. Texture analysis can be easily applied to routinely acquired MRI examinations. Therefore, retrospective studies (such as the current study) can be pursued to

Table 2 | Statistics comparing the traditional prediction model with the texture model

Statistic	Traditional model[a]	+Texture model
Pearson's r[b]	−0.51	−0.70
SE	3.9	2.6
Sres	21.6	17.9
Slope SE	0.155	0.094
Slope CI	[−1.307 to −0.693]	[−1.186 to −0.814]
Bias SE	3.95	2.64
Bias CI	[−7.81 to 7.81]	[−5.22 to 5.22]

eGFR, estimated glomerular filtration rate; HtTKV, height-adjusted total kidney volume; Sres, residual SD.
[a]Age, eGFR, and HtTKV.
[b]The square root of the adjusted r^2 for the prediction models.
The multiple linear regression models have a slope of −1 and bias of 0.

search for new imaging biomarkers of the disease when known disease progression data are available. Quantitative MRI techniques are more difficult to acquire and may be harder to standardize. So, although newer imaging techniques hold a great deal of promise in terms of tissue characterization and functional evaluation, their association with disease severity and disease progression will take many years to assess.

The phenotypical description provided by image texture features as they relate to characterizing the PKD kidney is important to understand in order to best apply these techniques in future studies. Briefly, we highlight these details for the 3 image texture features identified by stability selection and used in the regression model. *Entropy* measures the degree of disorder within the kidney. Kidneys with seemingly random cyst distributions will have a higher entropy measurement. *Correlation* measures grayscale value dependence of kidney voxels (the joint probability occurrence of specified pixel pairs). High values mean the kidneys are composed of many similar-appearing regions. *Energy* is the sum of squared elements in the gray level co-occurrence matrix (GLCM) and is a measure of tissue uniformity. The power of texture analysis comes from the fact that we can derive a number of measureable parameters regarding kidney structure that can be used to further refine an individual's prognosis based on phenotype.

The approach of texture feature analysis may also be preferable to other image processing techniques such as cyst segmentation. If cyst segmentation is performed manually, significant time is required, making a study of even moderate size nearly impossible. In addition, variability in the measurements will likely be fairly high, because discerning what are and are not cysts, as well as defining cyst edges, in scans of ADPKD patients is extremely difficult. On the other hand, if performed automatically, questions regarding the accuracy are difficult to ascertain. Fortunately, texture feature analysis provides information regarding cystic phenotype without requiring cysts to be individually segmented and classified.

CKD, which is characterized by reduced renal function, is a common outcome for ADPKD patients. However, the time course of disease progression to various stages of CKD is far from fully understood. Therefore, the development of new tools to better characterize the evolution of the disease is needed. We believe that the new imaging biomarkers provided by image texture feature analysis, along with automated TKV, will significantly improve the assessment of patient prognosis and will facilitate the ability to more quickly judge the effectiveness of interventions.

There are many challenges in terms of applying texture analysis to clinical practice as well as research studies. For example, image acquisition protocols, as well as image quality, should be kept as similar as possible because MRI shows a greater nonlinear influence on signal intensity and on quantification of heterogeneity (compared with CT). However, non–contrast-enhanced CT will not provide sufficient soft tissue contrast to make texture analysis effective. To be most effectively utilized, standardization of the image acquisition protocols is needed.

Future studies should be performed to evaluate texture analysis in larger and more diverse patient populations to fully understand how these new imaging biomarkers improve our understanding of ADPKD phenotypes for disease prognosis and clinical decision making. Adding other clinical data in nonbiased populations needs to be explored, as well as texture feature analysis on other imaging modalities (e.g., CT) and different sequences (e.g., T1-weighted MRI).

METHODS

Approval from Mayo Clinic's institutional review board was obtained for this retrospective study.

MRI data

The MRIs were coronal single-shot fast-spin echo (SSFSE) T2 sequences, acquired with a GE Medical Systems (Waukesha, WI) scanner, with matrix size 256 × 256 × Z (with Z large enough to cover the full extent of the kidneys within the imaged volume). Specifically, B_0: 1.5T; TE: 190 ms; TR: minimum; pixel size: 1.5 mm; slice thickness: 3.0 mm; slice spacing: 3.0 mm. Images were typically acquired over several breath holds.

Pre-processing

T2-weighted abdominal MRIs often have signal intensity nonuniformities, including both intra-slice[42] and inter-slice[43] artifacts. Intra-slice intensity artifacts come from several sources (e.g., choice of radio-frequency coil or sample geometry) that degrade image quality and introduce low-frequency intensity changes. Inter-slice intensity variations are the result of gradient eddy currents and cross-talk between slices. These cause interleaved "dark" and "bright" slices, which can be drastically different. To correct for these issues, we used an approach developed in-house that sequentially corrects inter-slice and intra-slice intensity variations.[44,45]

Finally, although texture analysis is rather robust to differences in absolute image signal ranges, the images were normalized based on the CSF in the adjacent spinal canal to have standardized tissue signal. Measurement of CSF intensity was performed for each scan by drawing a region-of-interest containing purely CSF. The mean value of this region-of-interest was then normalized to have a value of 1000 (i.e., CSF was normalized to have the same value in each MRI scan).

Kidney segmentation

A trained medical image analyst (MEE) performed kidney segmentation semiautomatically utilizing the MIROS software package,[46] which has an interactive viewer that allows visualization of the image data in coronal, sagittal, and axial planes. Segmentations can be overlaid and edited with a range of interactive tools. The MIROS algorithm was implemented in the Python programming language and has a push-button that starts the interactive tool for defining crude polygon contours of the cystic organ. We performed the MIROS method obtaining user input every third slice (every 9 mm). After segmentation of each kidney by MIROS, the interactive tools were used to perform quality assurance and finalize the segmentation.

Texture features

Gray level (first order). First-order texture features capture basic characteristics of the intensities within an image. Global intensity features were computed for each MRI. In addition, gradient and entropy local intensity features were computed at 5 different window sizes (7, 15, 23, 31, and 41). In order to account for the anisotropic nature of the voxel size, texture feature filters were calculated in a slice-by-slice fashion, for every slice within the volume.

Gray level co-occurrence matrix (second order). GLCM characterizes the second-order statistics of the spatial distribution of gray levels in an image.[47] For each image, 4 GLCM (corresponding to 4 directions) were calculated at 3 different window sizes (7, 15, and 31), and 6 features (contrast, dissimilarity, homogeneity, energy, correlation, and angular second moment) were derived from the GLCMs. The GLCM-based features were implemented in the Python programming language.

Texture-filtered features. Texture-filtered versions of the images were created for each patient. In the case of the second-order features, averaging over the 4 directions was performed (making the derived textures direction-invariant). Finally, for each of the kidney regions-of-interest, the mean, median, SD, skewness, and the 25th and 75th percentiles were calculated for the filtered textures. These values were then used as features to search for notable textural differences between the 2 groups of patients. Calculation of all texture features was performed in a matter of minutes for each patient scan.

Reproducibility. In order to determine the reproducibility of the image-derived texture features, 30 cases were reanalyzed by a second observer. Similarity metrics were calculated between the segmentations derived by the 2 observers, as well as comparison of the measured CSF values. Finally, the difference between the computed texture values was characterized.

Classification system

Feature selection. A robust feature selection method known as stability selection was used to identify candidate features for developing the multiple linear regression model.[48] The feature selection algorithm is applied on different subsets of data with different subsets of features. The process is repeated, and the number of times a particular feature is selected as important is evaluated. The most stable features get a score of 1, while weak features get a score of 0.

Multiple linear regression. The multiple linear regression modeling approach was implemented in Python with the sklearn toolkit, which uses a least-squares approach (function: sklearn.linear_model.LinearRegression). The approach minimizes the sum of the squared residuals in order to adjust the variable weights to best fit the data.

Statistical analysis

Analysis based on Wilcoxon rank-sum tests was performed to calculate the statistical significance of each imaging biomarker, as well as the area under the receiver operating characteristic curve (A_z), and the sensitivity and specificity of the features was also computed. Bootstrap resampling was used to estimate confidence intervals for A_z. Regression analysis was performed to characterize the relationship between individual biomarkers and the percentage change in eGFR.

DISCLOSURE

All the authors declared no competing interests.

ACKNOWLEDGMENTS

This work was supported by the National Institute of Diabetes and Digestive and Kidney Diseases under NIH Grant/Award Number P30 DK090728 to the Mayo Clinic Robert M. and Billie Kelley Pirnie Translational Polycystic Kidney Disease Center, the PKD Foundation under grant 206g16a, and the National Cancer Institute under grant/award CA160045. The CRISP study is supported by cooperative agreements from the NIDDK of the NIH (DK056943, DK056956, DK056957, DK056961) and by the National Center for Research Resources General Clinical Research Centers at each institution (RR000039, Emory University; RR00585, Mayo College of Medicine; RR23940, Kansas University Medical Center; RR000032, University of Alabama at Birmingham) and the National Center for Research Resources Clinical and Translational Science Awards at each institution (RR025008, Emory University; RR024150, Mayo College of Medicine; RR033179, Kansas University Medical Center; RR025777 and UL1 TR000165, University of Alabama at Birmingham; RR024153, University of Pittsburgh School of Medicine).

REFERENCES

1. Chapman AB, Devuyst O, Eckardt KU, et al. Autosomal-dominant polycystic kidney disease (ADPKD): executive summary from a Kidney Disease: Improving Global Outcomes (KDIGO) Controversies Conference. *Kidney Int*. 2015;88:17–27.
2. Grantham JJ, Torres VE, Chapman AB, et al. Volume progression in polycystic kidney disease. *N Engl J Med*. 2016;354:2122–2130.
3. Chapman AB, Bost JE, Torres VE, et al. Kidney volume and functional outcomes in autosomal dominant polycystic kidney disease. *Clin J Am Soc Nephrol*. 2012;7:479–486.
4. Bhutani H, Smith V, Rahbari-Oskoui F, et al. A comparison of ultrasound and magnetic resonance imaging shows that kidney length predicts chronic kidney disease in autosomal dominant polycystic kidney disease. *Kidney Int*. 2015;88:146–151.
5. Torres VE, Chapman AB, Devuyst O, et al. Tolvaptan in patients with autosomal dominant polycystic kidney disease. *N Engl J Med*. 2012;367:2407–2418.
6. Caroli A, Perico N, Perna A, et al. Effect of longacting somatostatin analogue on kidney and cyst growth in autosomal dominant polycystic kidney disease (ALADIN): a randomised, placebo-controlled, multicentre trial. *Lancet*. 2013;382:1485–1495.
7. Schrier RW, Abebe KZ, Perrone RD, et al. Blood pressure in early autosomal dominant polycystic kidney disease. *N Engl J Med*. 2014;371:2255–2266.
8. Mignani R, Corsi C, De Marco M, et al. Assessment of kidney volume in polycystic kidney disease using magnetic resonance imaging without contrast medium. *Am J Nephrol*. 2011;33:176–184.
9. Cohen BA, Barash I, Kim DC, et al. Intraobserver and interobserver variability of renal volume measurements in polycystic kidney disease using a semiautomated MR segmentation algorithm. *AJR Am J Roentgenol*. 2012;199:387–393.

10. Kline TL, Korfiatis P, Edwards ME, et al. Automatic total kidney volume measurement on follow-up MRIs to faciliate monitoring of ADPKD progression. *Nephrol Dial Transplant*. 2016;31:241–248.
11. Turco D, Severi S, Mignani R, et al. Reliability of Total Renal Volume Computation in Polycystic Kidney Disease From Magnetic Resonance Imaging. *Acad Radiol*. 2015;22:1376–1384.
12. Kistler AD, Poster D, Krauer F, et al. Increases in kidney volume in autosomal dominant polycystic kidney disease can be detected within 6 months. *Kidney Int*. 2009;75:235–241.
13. King BF, Reed JE, Bergstralh EJ, et al. Quantification and longitudinal trends of kidney, renal cyst, and renal parenchyma volumes in autosomal dominant polycystic kidney disease. *J Am Soc Nephrol*. 2000;11: 1505–1511.
14. Bae KT, Zhu F, Chapman AB, et al. Magnetic resonance imaging evaluation of hepatic cysts in early autosomal-dominant polycystic kidney disease: the Consortium for Radiologic Imaging Studies of Polycystic Kidney Disease cohort. *Clin J Am Soc Nephrol*. 2006;1:64–69.
15. Bae K, Park B, Sun H, et al. Segmentation of individual renal cysts from MR images in patients with autosomal dominant polycystic kidney disease. *Clin J Am Soc Nephrol*. 2013;8:1089–1097.
16. Chapman AB, Guay-Woodford LM, Grantham JJ, et al. Renal structure in early autosomal-dominant polycystic kidney disease (ADPKD): The Consortium for Radiologic Imaging Studies of Polycystic Kidney Disease (CRISP) cohort. *Kidney Int*. 2003;64:1035–1045.
17. Zhang JL, Morrell G, Rusinek H, et al. New magnetic resonance imaging methods in nephrology. *Kidney Int*. 2014;85:768–778.
18. Kajander S, Kallio T, Alanen A, et al. Imaging end-stage kidney disease in adults. Low-field MR imaging with magnetization transfer vs. ultrasonography. *Acta Radiol*. 2000;41:357–360.
19. Ebrahimi B, Macura SI, Knudsen BE, et al. Fibrosis detection in renal artery stenosis mouse model using magnetization transfer MRI. *Proc SPIE*. 2013;8672:867.
20. Kline TL, Irazabal MV, Ebrahimi B, et al. Utilizing magnetization transfer imaging to investigate tissue remodeling in a murine model of autosomal dominant polycystic kidney disease. *Magn Reson Med*. 2016;75:1466–1473.
21. Sourbron SP, Michaely HJ, Reiser MF, et al. MRI-measurement of perfusion and glomerular filtration in the human kidney with a separable compartment model. *Invest Radiol*. 2008;43:40–48.
22. Li LP, Halter S, Prasad PV. Blood oxygen level-dependent MR imaging of the kidneys. *Magn Reson Imaging Clin N Am*. 2008;16:613–625, viii.
23. Pedersen M, Dissing TH, Morkenborg J, et al. Validation of quantitative BOLD MRI measurements in kidney: application to unilateral ureteral obstruction. *Kidney Int*. 2005;67:2305–2312.
24. Prasad PV, Edelman RR, Epstein FH. Noninvasive evaluation of intrarenal oxygenation with BOLD MRI. *Circulation*. 1996;94:3271–3275.
25. Khatir DS, Pedersen M, Jespersen B, et al. Reproducibility of MRI renal artery blood flow and BOLD measurements in patients with chronic kidney disease and healthy controls. *J Magn Reson Imaging*. 2014;40: 1091–1098.
26. King BF, Torres VE, Brummer ME, et al. Magnetic resonance measurements of renal blood flow as a marker of disease severity in autosomal-dominant polycystic kidney disease. *Kidney Int*. 2003;64: 2214–2221.
27. Torres VE, King BF, Chapman AB, et al. Magnetic resonance measurements of renal blood flow and disease progression in autosomal dominant polycystic kidney disease. *Clin J Am Soc Nephrol*. 2007;2: 112–120.
28. Warner L, Yin M, Glaser KJ, et al. Noninvasive In vivo assessment of renal tissue elasticity during graded renal ischemia using MR elastography. *Invest Radiol*. 2011;46:509–514.
29. Maril N, Margalit R, Mispelter J, et al. Functional sodium magnetic resonance imaging of the intact rat kidney. *Kidney Int*. 2004;65:927–935.
30. Thoeny HC, De Keyzer F, Oyen RH, et al. Diffusion-weighted MR imaging of kidneys in healthy volunteers and patients with parenchymal diseases: initial experience. *Radiology*. 2005;235:911–917.
31. Chandarana H, Lee VS, Hecht E, et al. Comparison of biexponential and monoexponential model of diffusion weighted imaging in evaluation of renal lesions: preliminary experience. *Invest Radiol*. 2011;46:285–291.
32. Wang F, Kopylov D, Zu Z, et al. Mapping murine diabetic kidney disease using chemical exchange saturation transfer MRI. *Magn Reson Med*. 2016;76:1531–1541.
33. Castellano G, Bonilha L, Li LM, et al. Texture analysis of medical images. *Clin Radiol*. 2004;59:1061–1069.
34. Xie J, Jiang Y, Tsui HT. Segmentation of kidney from ultrasound images based on texture and shape priors. *IEEE Trans Med Imaging*. 2005;24: 45–57.
35. Lin DT, Lei CC, Hung SW. Computer-aided kidney segmentation on abdominal CT images. *IEEE Trans Inf Technol Biomed*. 2006;10:59–65.
36. Chen W, Giger ML, Li H, et al. Volumetric texture analysis of breast lesions on contrast-enhanced magnetic resonance images. *Magn Reson Med*. 2007;58:562–571.
37. Zhang J, Tong L, Wang L, et al. Texture analysis of multiple sclerosis: a comparative study. *Magn Reson Imaging*. 2008;26:1160–1166.
38. Yoshida H, Casalino DD, Keserci B, et al. Wavelet-packet-based texture analysis for differentiation between benign and malignant liver tumours in ultrasound images. *Phys Med Biol*. 2003;48:3735–3753.
39. Irazabal MV, Rangel LJ, Bergstralh EJ, et al. Imaging classification of autosomal dominant polycystic kidney disease: a simple model for selecting patients for clinical trials. *J Am Soc Nephrol*. 2015;26:160–172.
40. Levey AS, Inker LA, Matsushita K, et al. GFR decline as an end point for clinical trials in CKD: a scientific workshop sponsored by the National Kidney Foundation and the US Food and Drug Administration. *Am J Kidney Dis*. 2014;64:821–835.
41. Cornec-Le Gall E, Audrezet MP, Rousseau A, et al. The PROPKD Score: A New Algorithm to Predict Renal Survival in Autosomal Dominant Polycystic Kidney Disease. *J Am Soc Nephrol*. 2016;27:942–951.
42. Belaroussi B, Milles J, Carme S, et al. Intensity non-uniformity correction in MRI: existing methods and their validation. *Med Image Anal*. 2006;10: 234–246.
43. Schmidt M. *A Method for Standardizing MR Intensities Between Slices and Volumes*. Edmonton. Edmonton, AB, Canada: Department of Computing Science, University of Alberta; 2005.
44. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging*. 1998;17:87–97.
45. Tustison NJ, Avants BB, Gee JC. Directly manipulated free-form deformation image registration. *IEEE Trans Image Process*. 2009;18: 624–635.
46. Kline TL, Edwards ME, Korfiatis P, et al. Semi-automated segmentation of polycystic kidneys in T2-weighted magnetic resonance images. *Am J Roentgenol*. 2016;207:605–613.
47. Haralick RM, Shanmugam K, Dinstein IH. Textural features for image classification. *IEEE Trans Syst Man Cybern 3*. 1973;SMC-3:610–621.
48. Meinshausen N, Buhlmann P. Stability selection. *J R Stat Soc B*. 2010;72: 417–473.